DOCUMENT RESUME

ED 142 936 CS 003 537

AUTHOR Crawley, Sharon J.
TITLE Aural Cloze: A Review of Literature.
PUB DATE Apr 77
NOTE 24p.; Report prepared at University of Houston

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Cloze Procedure; *Language Research; *Listening Comprehension; *Listening Tests; *Literature Reviews

ABSTRACT

This paper reviews studies of the use of cloze techniques to measure listening comprehension. After briefly discussing the origins and characteristics of the cloze procedure, nine specific studies of the aural cloze are discussed. General conclusions are drawn from these studies and directions for future research are offered. (AA)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED142936

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Aural Cloze: A Review of Literature

Sharon J. Crawley

University of Houston

Houston, Texas

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Sharon J. Crawley

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND THE ERIC SYSTEM CONTRACTORS."

April 1977

CS003537

Abstract

Literature in which the cloze procedure has been used as a measurement of listening comprehension is reviewed in this paper. Secondly, suggestions for further research are offered.

Aural Cloze: A Review of Literature

Wilson Taylor is credited with developing the cloze procedure, which he first presented at an Association for Education in Journalism convention workship in 1953. Soon after, a description of the method appeared in the Journalism Quarterly (Taylor, 1953).

Originally, cloze was intended as a new approach to "readability." It derived its name from the term "closure" which refers to the tendency of human beings to complete a pattern which is familiar, but not-quite-finished. In language, people will complete an incomplete sentence pattern, "It is raining _____ and dogs." Most people, reading or hearing this sentence, would fill the blank in with the word "cats."

Taylor (1953) defines cloze as, "any single occurance of a successful attempt to reproduce accurately a part deleted from a 'message' (any language produce) by deciding, from the context that remains, what the missing word should be" (p. 416).

Taylor (1953) further defines the cloze procedure as:

> a method of intercepting a message from a 'transmitter' (writer or speaker), mutilating its language patterns by deleting parts, and so administering it to

> receivers (readers and listeners) that their attempts to make the patterns whole again potentially yield a considerable number of cloze units (p. 416).

In other words, the cloze procedure involves the completing of passages from which words have been systematically deleted.

Taylor (1956, 1953) and Bormuth (1968) describe the cloze readability procedure. Basically, the procedure involves selecting passages whose difficulty is to be determined, deleting every fifth word in the passage, replacing the deleted words with blanks of equal length, duplicating the tests, administering the tests to students who have not read them before, having students complete the blanks by filling in the word they think was deleted, and scoring the responses. The student has made a successful cloze if he completes the blank with the correct word.

On the surface, the cloze procedure may appear to be the same as a sentence completion or "fill-in-the-blank" test; however, if one looks more critically, one will notice differences. In sentence completion tests, the words are omitted after being pre-evaluated. In cloze, words are deleted systematically; there is no pre-evaluation involved. In sentence completion, whole phrases or clauses may be deleted. In cloze, only one word is deleted at a time.

The cloze procedure has been used in a variety of ways.

These uses include its: (a) correlation with standardized reading tests, (b) use as a measure of readability, (c) use as a measure of verbal ability, (d) use as a measurement of foreign language achievement, (e) use as a teaching device, and (f) use as a measure of listening ability.

Literature in which the cloze procedure has been used as a measurement of listening comprehension will be reviewed in this paper. Secondly, suggestions for further research will be offered.

Taylor (1956) reports that the cloze procedure can be used to discriminate between the "listenabilities" (p. 47) of materials. Two radio news stories of about 170 words each were read in two ways by an experienced radio announcer. The first version was written and read in a manner that was easy to comprehend; the second, in a manner that was difficult to comprehend. (Both versions followed grammatical rules.) The deletions were fourteen spoken words apart. A buzzer was sounded for about four seconds and a serial number was read for the deleted words. Subjects had to find the serial number on their answer sheets and write down their guesses for the missing words. Taylor states that the cloze scores discriminated between the "listenabilities" of "good" and "bad" script; the good script being more listenable.

Weaver (Dissertation Abstract, 1961) used oral and written presentations of cloze with two groups of juniors and seniors (80 subjects in each) at the University of Georgia to determine if there was a difference in the ability of subjects to complete the cloze. One group received reading tests and was allowed as many repetitions of the aural cloze as desired. The other group received the reading cloze and was allowed only one repetition of the aural cloze selection. Using a $2x2x^4$ analysis of variance, Weaver concluded that students performed significantly better in "any-word" silent cloze than aural, but that lexical deletion of nouns and main verbs was completed equally well by listening and silent reading.

Weaver and Kingston (1963) studied the relationship of cloze tests to standardized reading tests. One hundred sixty students in the Junior class at the University of Georgia took part in this study. Eight cloze tests were administered. These included a structural deletion of essay materials read silently, a lexical deletion of essay materials read silently, a structural deletion of a speech read silently, a lexical deletion of a speech read silently, a lexical deletion of essay material listened to, a structural deletion of a speech listened to, and a lexical deletion of a speech listened to.

These eight cloze tests were presented, in random order, to groups of eight to twelve students. Each test contained forty cloze items and the exact word was required to make a successful cloze. A thirty minute time limit was placed on the reading and listening cloze tests. Students were allowed to hear the listening tape through completely and to ask the examiners to reread any sentences in which a cloze item appeared before completing the cloze. Pauses of four seconds were used to indicate a deletion.

Other tests administered included the: Davis Reading Test, Modern Language Aptitude Test (MLAT) number learning subtest, MLAT phonetic script subtest, MLAT spelling subtest, MLAT words in sentences subtest, MLAT paired associates subtest, Sequential Tests of Educational Progress (STEP) listening subtest, Ohio State Psychological Examination (OSPE) vocabulary subtest, and OSPE reading comprehension subtest.

The correlations between the listening cloze tests and other tests ranged from .15 to .69. The highest correlation was .69 in which a structural deletion of a speech-listening was correlated with a lexical deletion of essay material-reading. The point needs to be made that these are both cloze tests. A correlation of .60 was found between lexical deletion of speech-listening and the Davis

Reading Test. A correlation of .61 was found between structural deletion of speech-listening and the Ohio State Psychological Examination (reading comprehension). Other positive correlations were found between structural deletion of essay material-listening and lexical deletion of essay material-reading; lexical deletion of a speech-reading and structural deletion of essay material-listening; and the STEP listening subtest and the listening cloze tests (r=.48, .45, .50, .54).

Jongsma (undated) wrote that the relatively low correlations "seem to suggest that the cloze listening tests were measuring factors other than listening comprehension" (p. 21).

Neville and Pugh (1971) compared the performance of children on similar reading and listening cloze tests with regard to the number and types of errors made. Sixty-six pupils (40 girls, 26 boys) ages eight years ten months to ten years nine months, from lower and lower middle class homes, and attending a Roman Catholic primary school in Yorkshire, were subjects in the experiment. The students were assigned randomly to two groups (X=32 subjects, Y=34 subjects). Forms R and B of the GAP Reading Test were used. (This is a cloze type test with approximately every tenth word deleted, and standardized in the United Kingdom). Both

forms of the GAP were recorded by a female adult reader with approximately ten second pauses for each deleted word. An answer paper with numbered lines for the deletions was prepared.

The two groups first received the reading versions of the GAP Test following manual instructions. After a short rest, groups X and Y received the listening version. In scoring, synonyms were not allowed, and errors were classified as omissions, wholly inappropriate responses, or incorrect responses according to the manual, but linguistically appropriate according to context.

The two means of the total sample for reading and listening were compared using a t-test for related (Total Reading M=24.60, Total Listening M=13.87, $t=5.93$, $p<.001$). A lower listening mean for the whole sample is in evidence. The listening mean average is equivalent to a reading age of eight years nine months on the GAP, the reading mean of ten years seven months, and the mean chronological age was nine years nine months. The results also pointed to a very high frequency of omissions on both forms of the listening version of the GAP.

Using the Pearson product moment, significant correlations ($r=.50$, $p<.01$) were found between the NFER Reading Test AD (which was administered one month before the investigation)

and the GAP Listening Test, and between the GAP Reading and GAP Listening Tests (r=.48, p .01). The results appear to establish GAP as a listening comprehension test.

Swalm (Dissertation Abstract, 1971) used the cloze procedure (with a ten percent random deletion of lexical words) with 324 students in grades two, three and four (108 from each) to study the effect of testing materials that had readabilities equivalent to students' grade placements upon oral and silent reading and listening comprehension.

Using analysis of variance formulas and Tukey's Honestly Significant Difference Test (significance established at $p < .05$) oral reading was significantly better than silent reading or comprehension or listening comprehension at the second grade level; no significant differences were found between methods at the third and fourth grade levels; silent reading of above-average students was significantly better than listening in grades two and four; and students with average reading ability tended to comprehend equally well with all three approaches (significance not found). Listening was significantly more effective for comprehension than silent reading for below-average students at all three grade levels, and significantly different from oral reading only in grade three.

Kennedy and Weener (1973) used the cloze procedure as a

means of, and to determine if, training poor readers to attend auditorially and visually to contextual cues would improve listening and/or reading comprehension. More specifically, Kennedy and Weener hypothesized that: (1) experimental groups receiving visual and auditory training would perform significantly better than control groups on listening and reading tests; and (2) students who received auditory training would make the greatest gains on listening tests; and students who received visual training would make the greatest gains on reading comprehension tests.

Eighty students, completing third grade, with reading grade levels between 1.9 and 3.5 (as determined by the Metropolitan Achievement Test) and reading between the 1-2 (first grade, second half) and 3-1 (third grade, first half) on paragraphs from the Open Court Basic Readers were used in the study. The subjects were ranked from high to low according to test results, and assigned to one of four groups. The top four students were assigned randomly to one of the four groups, the next four were assigned randomly, etc., until all eighty students were assigned to a group (20 students in each group; two experimental groups and two control groups).

The experimental listening group received five, 20-minute, training sessions (a total of one hour forty minutes) in replacing words deleted from sentences and stories (ten

sentences and one story during each session) which were presented auditorially. A bell was rung each time a word was deleted, and subjects could have sentences repeated. The experimental reading group received similar training, but the sentences and stories were presented visually. To account for the possible Hawthorne effect, the control reading group received individualized oral reading instruction for five, twenty minute, training sessions. The control class group remained in the regular classroom and received no special treatment.

The cloze materials were designed with one word deleted from each sentence; and an attempt was made to have words taken proportionally from the beginning, middle and end of sentences. Also, prepositions and conjunctions, verbs and adverbs, adjectives and articles, and nouns were taken out during one session each. A successful cloze was made if the word was "syntactically or semantically correct" (p. 533).

Using a Least Significant Difference test it was found that: (1) the experimental listening group scored higher on the Durrell Listening Reading Series listening comprehension subtest than the reading comprehension subtest ($p < .10$); (2) the experimental reading group scored significantly higher ($p < .05$) on the reading comprehension subtest of the DLR than on the listening comprehension subtest;

(3) the experimental groups did significantly better ($p < .05$) than the control class on the listening comprehension subtest; (4) the experimental groups performed significantly better ($p < .05$) than the two control groups on the Cloze Procedure subtests; and (5) there was no significant difference between the experimental groups on the Cloze Procedure tests.

Craker (Dissertation Abstract, 1971) used the Clozentropy Procedure to test the oral English ability of 100 pre-literate lower middle class Angle, Black, Navajo and Spanish children (25 in a group) of legal age for first grade. An Oral Clozentropy Test of English Proficiency was developed, and a CDC 6400 computer program analysis of variance was conducted to compare group means. Craker concluded that 98 of the 100 students understood and could use the procedure.

Smith (Dissertation Abstract, 1974) studied the ability of 84 kindergarten children to use context on aural cloze. She investigated which parts of speech (nouns, verbs, adjectives, and adverbs) were easiest to supply, and the easiest sentence position for children to complete the cloze. Tests were administered individually and orally to each student and exact and equally acceptable replacements were allowed in scoring. Using a one-way analysis of variance and multiple comparison procedures according to the Scheffé

method, she concluded that students found nouns in end position the easiest to complete and verbs in end position the next easiest. Noun deletions were found to be easiest and verb, the next easiest of random deletions. Cloze scores did not increase significantly as a result of accepting equally acceptable responses; and there was no sex difference in performance.

Kingston and Weaver (1970) attempted to determine the feasibility of using the cloze procedure as a teaching technique with disadvantaged rural children, and to determine its predictive power with achievement tests at the first grade level. One hundred and eighty-two rural, white disadvantaged first grade students were taught the cloze procedure. Approximately half of the students were used to obtain reliability data for the instruments, eight were dropped from the study because they were absent during testing or left school; data from 74 students was used for the regression analysis which was calculated.

A planned strategy was used to teach students how to use the cloze procedure. Using the language experience approach, the children made up stories; and the teacher wrote these on the board. The teacher then omitted a lexical item and the children supplied words they thought could complete the passage and still make sense. The next step involved typing

stories the students developed and presenting them to the students at a later time. Words were then deleted from these stories. At a more advanced level, students were shown books which had rebus pictures. The number of pictures presented was decreased gradually and blanks were used in place of them. Students supplied the missing words. In the next stage, the teacher read aloud stories which the students had written and the students followed along. Deleted versions of the stories were then read aloud by the teacher and students supplied the missing words. At first only two to three words were deleted per 100-200 word stories. As a final step, in the cloze procedure training, students were given stories they had not written in class, and asked to fill-in the missing words. Words which had the first letter correct in their spellings were counted as correct.

The Lee-Clark Readiness Test was administered during the week preceeding, and also during the first week of school. The Ginn Pre-Primer test was administered approximately mid way through the year. Cloze tests were administered during February and March. The Ginn Primer Test was administered during April, and the California Achievement Tests during May.

Four types of cloze tests were administered. The first type was an "any-word" cloze. With this test, students

read the passage containing all the words before completing the cloze test. Every fifth or tenth word was deleted on a 50 items cloze test which was administered in five sessions with ten items each. The test-retest reliability for this test was .78. The second cloze test was a "multiple-choice, structural" cloze. Each deleted word was paired with a distractor of the same grammatical class. After five deletions, ten words were listed across the page from which the student could select choices. The test-retest reliability was .91. The third type of cloze test was a "multiple-choice, lexical" cloze. A fifth word deletion was carried out on nouns, verbs and adjectives. Each word was paired with a distractor of the same grammatical class. After five deletions, ten words were listed across the page. The test-retest reliability was .76. The last type of cloze used was an "aural reading" cloze. This is the same as the "any word" cloze, except that the teacher reads the passage aloud as the student reads it silently. The teacher stops for thirty seconds at each cloze blank. The test-retest reliability was .85.

A step-wise regression was computed on the California Reading Test using the cloze tests, Lee Clark Readiness, Ginn Pre-Primer and Ginn Primer Tests as independent variables (Sig. F at .01 level for all). Aural-cloze correlation

coefficients with other testing measures were: California Reading .62; California Arithmetic .66; California Language .57; Multiple-choice, Structural Cloze .87; Multiple-choice, Lexical Cloze 173; and Straight Cloze .76.

Kingston and Weaver's best conclusion appears to be that first grade children were able to handle oral and written cloze. They also concluded that the cloze tasks were better predicotrs of standard reading scores obtained at the end of the year than were the readiness and basal reader instruments.

Little research has been conducted using the aural cloze procedure as a measure of listening comprehension. Nine studies have been cited in this paper. From these studies, the following conclusions appear to be in order.

Aural cloze may be used as a means of discriminating the listenabilities of materials. Taylor (1956) concluded that cloze scores discriminated between the "listenabilities" of "good" and "bad" script,--the good script being more listenable.

Aural cloze may be used as a measure of listening comprehension. Weaver (1961) determined that nouns and main verbs could be filled-in equally well by silent reading or listening. Weaver and Kingston (1963) found a

significant correlation ($p < .05$) between aural and reading cloze passages, aural cloze tests and standardized reading comprehension tests, and aural cloze tests and the STEP listening test. Nevill and Pugh (1971) found significant correlations between the NFER Reading Test and the GAP Listening Test ($p < .01$) and between the GAP Reading Test and the listening version ($p < .01$). Swalm (1971) found that students with average reading ability comprehended grade level oral and silent reading and listening material equally well when the cloze procedure was used as a testing instrument. Kennedy and Weener (1973) found that students trained in the cloze procedure scored significantly better ($p < .05$) on the Durrell Listening-Reading (listening comprehension subtest) than the control group; and there were no significant differences between the experimental groups on the Cloze Procedure Tests.

Noun deletions appear to be the easiest to supply in aural cloze. Smith (1974) concluded that noun deletions were the easiest to supply, followed by main verbs.

Aural cloze may be used with students in first grade. Cracker (1971) concluded that children of first grade age could understand and perform on aural cloze tests. Kingston and Weaver (1970) concluded that first grade children were able to handle oral and written cloze. And Smith (1974)

successfully used kindergarten students in a study of the use of context on aural cloze passages.

There is a tendency for researchers to attempt to ascertain the relationship between aural (listening) cloze tests and measures of silent reading rather than listening comprehension. Weaver and Kingston (1963) correlated cloze tests with the Davis Reading Test, Modern Language Aptitude Test and the Ohio State Psychological Examination. Kingston and Weaver (1970) correlated aural cloze tests with the California Reading Test, Lee Clark Readiness, Ginn Pre-Primer, and Ginn Primer Tests. Neville and Pugh (1971) correlated the reading form of the GAP Reading Test with a listening version of the GAP and the NFER Reading Test.

Further investigation into the relationship between aural cloze tests and standardized listening tests such as the Durrell-Listening Reading: Listening Subtest, Stanford Achievement: Listening Comprehension Subtest, Assessment of Children's Language Comprehension, Brown-Carlson Listening Comprehension Test, Cooperative Primary Tests: Listening, Orr Graham Listening Test, Progressive Achievement Tests of Listening Comprehension, Sequential Tests of Educational Progress: Listening, and Tests for Auditory Comprehension of Language appear in order.

Investigation into the relationship between aural cloze tests and the listening comprehension subtests of reading inventories such as the Denver Public Schools Reading Inventory, Classroom Reading Inventory by Silvaroli, Diagnostic Reading Scales by Spache, Standard Reading Inventory by McCracken, and the Individual Reading Placement Inventory (Follett Educational Corporation) appear in order.

Another area to investigate is whether sex differences exist in aural cloze results. Smith (1974) found that "there is no evidence to indicate that there are sex differences for...aural cloze tasks" (p. 3550-A) in her investigations with kindergarten children. Investigations at other grade levels are in order.

A fourth area to investigate is the predictive power of aural cloze in assessing reading readiness. Craker (1971) used an orally administered cloze test with children of first grade age. Smith (1974) showed that kindergarten children could use aural cloze. Kingston and Weaver (1970) showed that first grade children could use aural cloze. They also attempted to determine its predictive power over a readiness test with achievement tests, but administered the readiness test at the beginning of the school year and didn't administer the aural cloze tests until February and March, well into the school year.

A review of available literature on aural cloze and a number of areas which are in need of further investigation have been presented in this paper. It should be apparent that there is much to investigate in the use of aural cloze.

References

Bormuth, John R. Cloze test readability: Criterion reference scores. Journal of Educational Measurement, 1968, 5, 189-196.

Craker, Hazel V. Clozentropy procedures as an instrument for measuring oral English competencies of first grade children (Doctoral dissertation, University of New Mexico, 1971). Dissertation Abstracts International, 1972, 32, 6286A-6287A. (University Microfilms No. 72-13,781)

Jongsma, Eugene A. The cloze procedure: A survey of the literature. (Occasional Papers in Reading). Unpublished manuscript, Indiana University, undated.

Kennedy, Delores K. & Weener, Paul. Visual and auditory training with the cloze procedure to improve reading and listening comprehension. Reading Research Quarterly, 1973, 8, 524-541.

Kingston, Albert J. & Weaver, Wendell W. Readibility of cloze techniques for teaching and evaluating culturally disadvantaged beginning readers. Journal of Social Psychology, 1970, 82, 205-214.

Nevill, Mary H. & Pugh, A.K. Context in reading and listening: A comparison of children's errors in cloze tests.

British Journal of Educational Psychology, 1971, 44, 224-232.

Smith, Paula Jean. Kindergarten children's abilities to use context on aural cloze passages. (Doctoral dissertation, University of Arizona, 1974). Dissertation Abstracts International, 1974, 35, 3500-A. (University Microfilms No. 74-28,313)

Swalm, James E. Comparison of oral reading, silent reading, and listening comprehension assessed by cloze. (Doctoral dissertation, Rutgers University The State University of New Jersey, 1971). Dissertation Abstracts International, 1972, 32, 3578-A. (University Microfilms No. 72-1102,160)

Taylor, Wilson S. Cloze procedure: A new tool for measuring readability. Journalism Quarterly, 1953, 30, 415-433.

Taylor, Wilson S. Recent developments in the use of "cloze procedure." Journalism Quarterly, 1956, 33, 42-48 & 99.

Weaver, Wendell W. An examination of some differences in oral and written language using the cloze procedure. (Doctoral dissertation, University of Georgia, 1961). Dissertation Abstracts International, 1961, 22, 2702.

Weaver, Wendell W. & Kingston, Albert J. A factor analysis of the cloze procedure and other measures of reading and language ability. Journal of Communication, 1963, 13, 252-261.